PERGAMON

Vision Research 42 (2002) 2153–2162

Vision Research

www.elsevier.com/locate/visres

Contour completion through depth interferes with stereoacuity

Dawn Vreven *, Suzanne P. McKee, Preeti Verghese

Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115, USA

Received 12 November 2001; received in revised form 4 March 2002

Abstract

Local disparity signals must interact in visual cortex to represent boundaries and surfaces of three-dimensional (3D) objects. We investigated how disparity signals interact in 3D contours and in 3D surfaces generated from the contours. We compared flat (single disparity) stimuli with curved (multi-disparity) stimuli. We found no consistent differences in sensitivity to contours vs. surfaces; for equivalent amounts of disparity, however, observers were more sensitive to flat stimuli than curved stimuli. Poor depth sensitivity for curved stimuli cannot be explained by the larger range of disparities present in the curved surface, nor by disparity averaging, nor by poor sensitivity to the largest disparity in the stimulus. Surprisingly, sensitivity to surfaces curved in depth was improved by removing portions of the surface and thus removing disparity information. Stimulus configuration had a profound effect on stereo thresholds that cannot be accounted for by disparity-energy models of V1 processing. We suggest that higher-level 3D contour or 3D shape mechanisms are involved.

Keywords: Contour; Depth; Disparity; Filling-in; Psychophysics

1. Introduction

Stereoacuity is the ability to judge relative depth from disparity. It is thought to be limited by the sensitivity of disparity detectors in primary visual cortex (Farell, 1998; Morgan & Castet, 1997). Classic work (Badcock & Schor, 1985; Blakemore, 1970; Ogle, 1953; Westheimer & McKee, 1978) has measured stereoacuity between two lines or two gratings, each presented with a "standing" or pedestal disparity in addition to the small relative disparity between the lines. The resulting disparity increment function indicates that relative depth sensitivity (i.e., the inverse of disparity threshold) decreases as an exponential function of standing disparity. These data make good physiological sense, given the response properties of disparity-selective cells in V1. The decrease in stereoacuity would be expected if units tuned to lower spatial frequencies were responsible for coding larger disparities. Others have measured stereoacuity using sinusoidally modulated contours (Tyler, 1975) or modulated random-dot disparity gratings (Schumer & Julesz, 1984; Tyler, 1974), which appear as surfaces corrugated in depth when viewed stereoscopically. Schumer and Julesz (1984) added a standing disparity to disparity-modulated random-dot surfaces. They also reported an exponential rise in the amplitude of disparity required to discriminate a corrugated surface from a flat surface as standing disparity is increased. These outcomes suggest that similar neural mechanisms are used to combine disparity signals in flat surfaces and in surfaces curved through depth.

This paper approaches stereoacuity by examining sensitivity to three-dimensional form. Unlike previous investigations of stereoacuity, we will address the issues of boundary representation and surface filling-in. Stereoscopic contours that varied in depth were used to generate surfaces that varied in depth. Thus, depth at the center of the 3D surface was specified only at the left and right vertical contours (see Fig. 1b). Despite the complete absence of explicit disparity information in the region between these two vertical contours, observers perceive the entire surface in compelling depth. Our question is: how sensitive are observers to the depth in these surfaces? The question is important because it tells us about the nature of form representation both across space and through depth. Consider how these stimuli present the visual system with a two-part problem: first, the disparity signals along a contour must be integrated

* Corresponding author.
E-mail address: vreven@uwosh.edu (D. Vreven).

0042-6989/02/$ - see front matter
PII: S0042-6989(02)00126-8

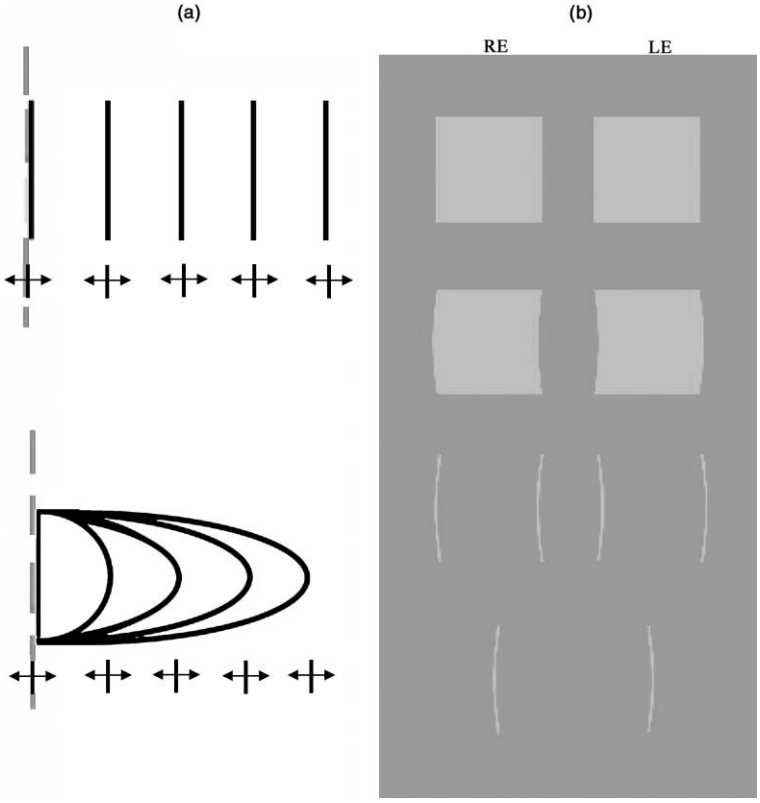

Fig. 1. (a) Profile of flat (top panel) and curved (bottom panel) stimuli. The dashed line represents the fixation plane. Each stimulus is shown with a depth probe beneath it. Five standing disparities are shown for flat stimuli, and five equivalent peak disparities are shown for curved stimuli. The surfaces with the largest standing/peak disparity (11.2 min) contained 7.5 cm of depth at the viewing distance used in the experiments. (b) Stereograms of experimental stimuli, arranged for cross-fusion. From top to bottom: flat surface, curved surface, two curved contours, one curved contour. RE: right eye, LE: left eye.

(Hess, Hayes, & Kingdom, 1997) to recover the depth boundaries of the surface. Second, these depth boundaries are presumably propagated to the center of the surface (Collett, 1985; Wurger & Landy, 1989). Our strategy was first to compare sensitivity for flat and curved surfaces in order to examine boundary recovery through depth. Next, we compared contour sensitivity and surface sensitivity to examine filling-in. To anticipate, we found no difference in disparity sensitivity between surfaces and contours. There was a large and reliable sensitivity difference, however, between flat and curved stimuli.

2. Method

2.1. Observers

Four observers participated. Each observer had normal or corrected to normal vision and stereoacuity to at least 20 arc s as measured by the StereoOptical RandDot circles. Two observers were two of the authors who were highly practiced at making stereoscopic depth judgments. The remaining observers were naïve about the purpose of the experiments and less practiced.

2.2. Stimuli and experimental procedure

Stimuli were generated using custom software and displayed on a pair of 15-in. CRT monitors, each with a P4 phosphor and a refresh rate of 71 Hz. Observers viewed stimuli in a dimly-lit room through a modified Wheatstone mirror stereoscope. A chin rest stabilized the head. Two sets of mirrors were fixed in position, resulting in a fused virtual image at a viewing distance of 122 cm (the distance to the virtual image was the same as the distance to either monitor). At this viewing distance, a single pixel subtended 7 min arc. Each eye received a half-image from one of the monitors, which were calibrated for both congruent visual direction and luminance.

Test stimuli were half-images of flat (single-disparity) and curved (multi-disparity) stereoscopic contours and surfaces. Surface and contour luminance, measured by a

Pritchard photometer on a 6 min by 6 min test patch, was 130 candelas/m^2 at 1 m distance; the background luminance was 45 candelas/m^2. Each contour was 3.5 min wide and extended 2.3 deg vertically. In general, each half-image contained two contours separated by 2.3 deg horizontally (Fig. 1). Surface stimuli were generated by filling the background-luminance portion between the two contours in each half-image with high luminance. Thus, surface stimuli subtended 2.3 deg^2 on the retina. The resulting percept was a textureless surface whose depth at any point was congruent with its left and right vertical contours. For some displays, only a single central contour was presented.

Standing disparity was added to flat (single-disparity) stimuli by shifting the half-images nasally relative to the fixation point. This manipulation generates crossed disparity at the vertical contours in the image. Five standing disparities were used, ranging from 0 to 11.2 min of disparity. Standing disparity was not added to curved (multi-disparity) stimuli; rather, the amount of curvature in depth was manipulated to be equal to the standing disparity of flat stimuli. This was achieved by curving the vertical contours according to a cosine function ($y = r \times \cos x$, where r is the peak of the surface or, equivalently, the point along the contour with maximum disparity). Two fixation conditions were used with curved surfaces: either the top and bottom horizontal edges of the surface appeared in the fixation plane, in which case the vertical contours were given crossed disparity, or the peak of the surface appeared in the fixation plane, in which case the vertical contours were given uncrossed disparity. In all conditions, five curvatures were used. The disparity at the peak of the curved surface was identical in magnitude to the standing disparity used for flat stimuli, and therefore ranged from 0 to 11.2 min.

A small probe line (length 0.5 deg, width 1.4 min) with the same luminance as the test stimuli appeared in each half-image, centered 0.5 deg below the bottom of the test stimulus. The probe line took on five disparity values, always centered on the disparity of either the flat stimulus or the peak of the curved stimulus. If one considers the probe's center disparity value as 0, then the probe was given two crossed and two uncrossed disparity values relative to this point. The disparity step size for the probe was variable depending on the observer, but typically the entire range of probe disparity was under 1 min. A luminance dithering technique was used to obtain these subpixel shifts in probe position.

At the beginning of each trial, a button press initiated the appearance of a fixation dot with two vertical nonius lines, one above and one below the dot. The observer was instructed to change vergence in order to align the nonius lines vertically before pressing a button to initiate the trial. Once initiated, the fixation and nonius lines disappeared and the test stimulus appeared centered 0.25 deg above the fixation point, with the line probe centered 0.25 deg below the fixation point. The stimulus duration was 200 ms, a duration too brief to complete voluntary vergence eye movements. The observer's task was to indicate whether the probe was in front of or behind the test stimulus. For curved stimuli, the observer indicated whether the line probe was in front of or behind the peak. Feedback was always given, and observers were required to practice until thresholds stabilized (ranging from 20 to several hundred trials depending on the observer). A probit function was fit to each 100 trial session, yielding estimates of both the threshold and the mean. Typically, threshold estimates are based on three or more sessions, and a measure of the error of the threshold estimate is given by the standard error (SE) of the threshold across 100-trial sessions. The means were examined for deviations from zero; because feedback was given and practice was required, the means were generally close to zero with no systematic deviations.

Finally, a jitter disparity was added to every display. The jitter was a random 2–5 pixel shift (either nasally or temporally) of all elements in the display, which added up to 3.5 min of disparity to the display. This was done to assure that observers were not using static depth cues, such as the frame of the monitor, to aid depth judgments.

3. Results

3.1. Experiment 1: surfaces

Disparity thresholds for flat and curved surfaces are shown in Fig. 2 as a function of standing/curvature disparity (white symbols). Consider first the flat surface (square symbols, dashed lines). Thresholds increase as standing disparity is increased, consistent with the classic work. The absolute value of the thresholds is slightly higher than reported by Blakemore (1970); this is probably due to the addition of whole field disparity jitter to ensure that observers were not using the monitor frame as a reference for depth judgments (see Section 2). Thresholds for curved surfaces (white circles), however, increase more rapidly with increasing curvature disparity. At disparities greater than about 6 min, curved surfaces have higher thresholds than flat surfaces. This is surprising, because the disparity about which the judgment is being made is the same for flat and curved surfaces, leading to the expectation of equal sensitivity.

One possible explanation for poor sensitivity with curved surfaces is that the peak of the surface is diplopic relative to the fixation plane at the back of the surface. To test this notion, we arranged for both the peak and

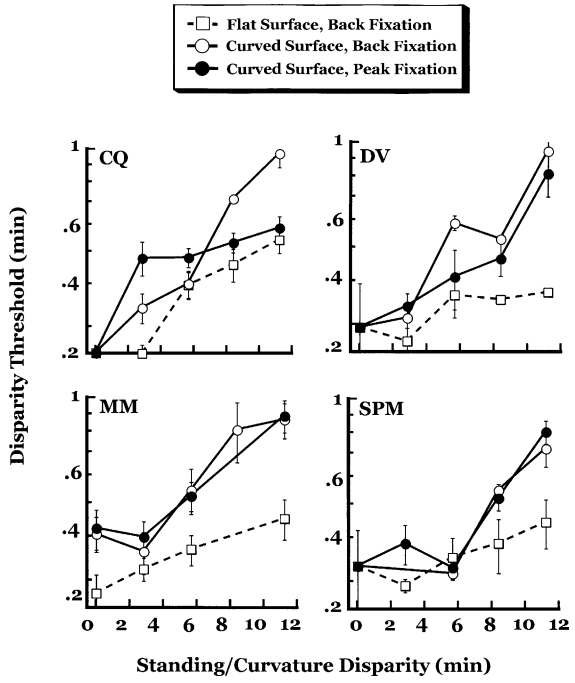

Fig. 2. Disparity thresholds for flat (white squares, dashed line) and curved (white circles) surfaces with equivalent disparity. Thresholds for flat surfaces increase with standing disparity, but thresholds for curved surfaces increase more rapidly, particularly as the disparity exceeds 6 min arc. Placing the peak of the curved surface (where the depth judgment is being made) in the fixation plane has little effect on performance for three of four observers (black circles).

the probe to appear in the fixation plane, where presumably the most sensitive disparity mechanisms could operate (Fig. 2, black circles). Three of four observers are either unaffected or show a slight improvement in performance when the peak of the surface is in the fixation plane. This improvement is not large enough to explain the difference in stereoacuity between flat and curved surfaces, but it is consistent with an inverse weighting of disparity signals as a function of depth from the fixation plane (Backus, Banks, van Ee, & Crowell, 1999). Apparently, observer CQ is strongly affected by distance from the fixation plane. To assure that this factor does not confound our results, we ran all remaining experiments with the peak of the curved surface in the fixation plane.

Additionally, we confirmed that thresholds were not dependent on monocular curvature by presenting a single half-image of the 11.2 min-curved surface stimulus and asking an observer (DV) to judge whether the probe appeared to the left or the right of the center of the surface. Thresholds were approximately 3.5 times higher with only monocular curvature (threshold = 2.86, SE = 0.09 monocularly vs. threshold = 0.81, SE = 0.11 binocularly). This outcome is consistent with previous studies showing that monocular signals are, in some cases, less precise than binocular signals (McKee, Levi, & Bowne, 1990a,b) and do not account for the sensitivity we observe. [1]

3.2. Experiment 2: contours

Next, we compared stereoacuity for contours to that for surfaces. For convenience, thresholds from the most curved surface in the peak fixation condition in Section 3.1 and thresholds from the flat surface with the greatest standing disparity in Section 3.1 are re-plotted in Fig. 3 (curved surface and flat surface conditions, respectively). Consider first the curved stimuli. Thresholds for the curved surface are indistinguishable from thresholds obtained when only the left and right contours were presented (two curved contours condition). Further, there are no systematic changes in threshold when only a single contour is presented in the middle of the display (one curved contour condition). Similarly, thresholds for a single flat contour were not different from flat surface thresholds (flat surface and flat contour conditions in Fig. 3). Thus, we find no consistent differences in sensitivity between the source of the disparity signal (contours) and the surface which is filled-in between the contours. This suggests that the limiting factor in sensitivity to surfaces is not filling-in per se.

These results also address the role of eccentricity on disparity sensitivity. One could argue that thresholds are high for curved surfaces because the relevant disparity signal at the peak of the surface is located at a greater eccentricity than the signal for flat surfaces. Note that both surfaces contain no disparity information along the bottom edge. This is because uniform, horizontally-oriented lines do not contain horizontal disparity except at the endpoints. However, the bottom left corner (say) of the binocularly-presented flat surface could be used to make a disparity judgment, compared with the midpoint of the left contour for the curved surface. The data in Fig. 3 show that decreasing the eccentricity of the contours does not affect performance, and therefore cannot account for the sensitivity difference between flat and curved surfaces. Likewise, the idea that the disparity of the probe interacts with the disparity of the contours is not supported, since contours can be presented directly above the probe or ~2 deg eccentric and there is no change in disparity sensitivity.

[1] There is no consistent relationship between stereoacuity and the monocular sensitivity for position changes. For some configurations, stereoacuity is greatly superior to position judgments, e.g, when the test and reference bars are widely separated and presented in the fixation plane (Berry, 1948; McKee et al., 1990b; Westheimer & McKee, 1979). For other configurations, position judgments are greatly superior to stereo judgments, e.g. spatial interval judgments along the x-axis are generally better than depth interval judgments along the z-axis for intervals greater than 2 arcmin (McKee et al., 1990a).

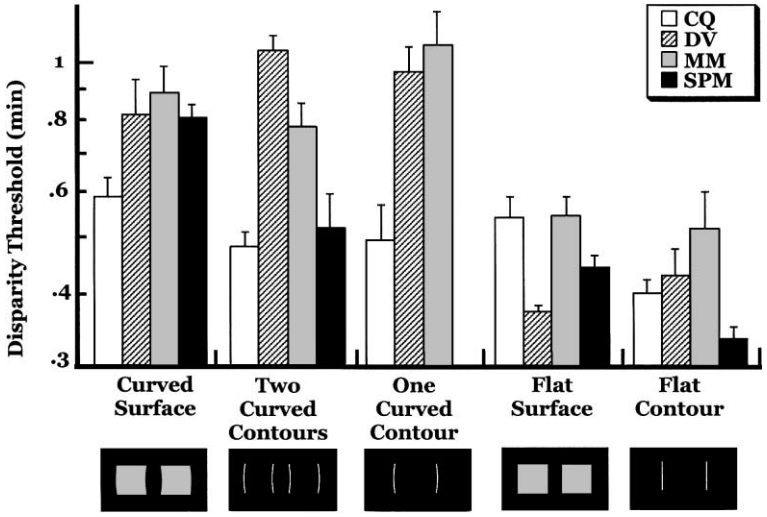

Fig. 3. Disparity thresholds for four observers showing that thresholds are higher for curved stimuli than for flat stimuli. Each condition is illustrated with a schematic stereogram below the absicca. Data are from conditions with the largest peak disparity for curved stimuli or the largest standing disparity for flat stimuli (11.2 min). For curved stimuli, the peak of the stimulus appeared in the fixation plane. The curved surface data are re-plotted from Section 3.1. For three of four observers, thresholds for curved stimuli are elevated relative to those for flat stimuli. The pattern holds regardless of whether the stimulus is a luminance surface, two eccentric luminance contours, or a single, centralized, luminance contour.

3.3. Experiment 3: random-dot surfaces

The large range of disparities in curved stimuli may account for threshold elevation relative to flat stimuli. Consider that the disparity gradient (change in disparity across spatial extent) for flat surfaces was zero, whereas the gradient for curved surfaces was both non-zero and variable across the surface. To determine if these factors could account for our results, we presented a dense (151 dots/deg^2) random-dot stereogram (RDS) matched to the most-curved luminance surface in apparent brightness and peak disparity. Thus, the RDS contained the same range of disparities and the same rate of disparity change as the luminance surface. Thresholds for the RDS were a factor of 2 lower on average than thresholds for the luminance surface (RDS and curved surface conditions in Fig. 4, respectively). This outcome argues against the notion that differences between flat and curved stimuli were due to the range of disparities or the steepness of disparity gradients present in the stimuli. Further, note that the only difference between the RDS and peak fixation conditions is whether the curved surface is generated with random dots or luminance contours. Curiously, it appears that performance is much worse for curved stimuli whose disparity is specified by luminance contours than it is for curved stimuli whose disparity is defined by random dots.

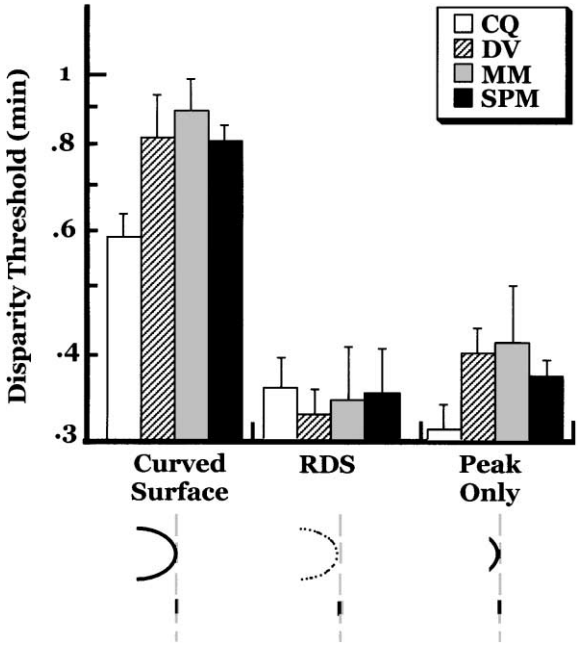

Fig. 4. Disparity thresholds for four observers showing that curvature per se does not account for elevated thresholds. Each condition is illustrated with a schematic profile of the stimulus below the absicca. The dashed line in the profile represents the fixation plane and black lines represent the test stimuli and probe. Data are from conditions with the largest curvature disparity (11.2 min). As in Fig. 3, the curved surface data are re-plotted from Section 3.1. When a random-dot surface is presented, thresholds improve dramatically (RDS). Thresholds are also low when only a small portion of the surface is presented (Peak Only).

3.4. Experiment 4: partial surfaces

One explanation for our results is that detectors responding to the multiple disparities in luminance stimuli interact with one another. This might be expected if spatially adjacent disparity detectors, arranged along the

disparate contour, underwent cooperative or competitive interactions to arrive at a local estimate of the disparity (Grossberg, 1997). To test this notion, we first created a "Peak Only" stimulus by assigning background luminance to all portions of the curved surface stimulus except a 0.5 deg (vertical) by 2.3 deg (horizontal) strip through the center of the surface. This left the curved peak of the surface intact. Thresholds again decreased to the level found for flat surfaces (Peak Only condition in Fig. 4). This is surprising because it suggests that, in some cases, sensitivity improves with less disparity information, supporting the notion of interaction between disparity-tuned units.

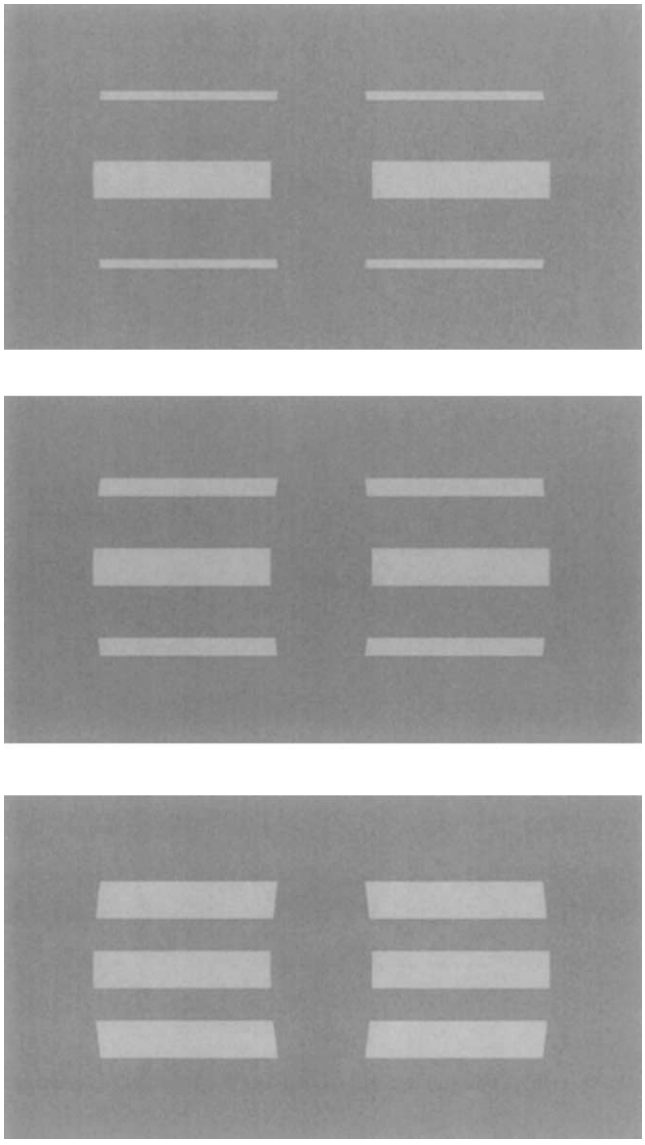

Fig. 5. Stereograms of the partial surface stimuli, arranged for cross-fusion. Note that each stimulus contains the peak of the surface as well as the top and bottom edges at the back of the surface. Thus, depth judgments in each case could be based on just the peak; further, each stimulus has the same range of disparity. The percent of surface area in each stimulus, from top to bottom, is 30%, 40%, and 60%.

If disparity detectors are subject to interaction from other units, then adding surface portions to the stimulus (and thus adding disparity) should increase thresholds. To test this hypothesis, a series of partial surface stimuli were created by adding high luminance 2.3-deg-wide "strips" of various heights to the Peak Only stimulus (Fig. 5). Strips were always added symmetrically from the top and bottom horizontal contours of the surface, resulting in stimuli whose area was approximately 30%, 40%, and 60% of the original. The strip heights were 0.12 deg, 0.23 deg, and 0.47 deg (from top to bottom in Fig. 5). The peak segment was left intact in each stimulus. Presumably, the peak segment alone could be used to make depth judgments, regardless of the additional high-luminance strips.

Fig. 6 plots disparity thresholds as a function of surface area for the most curved stimulus (disparity = 11.2 min). The Peak Only stimulus (20% surface area) and the intact stimulus (100% surface area) are re-plotted here for comparison. Adding more disparity (in the form of increasing the surface area) increased thresh-

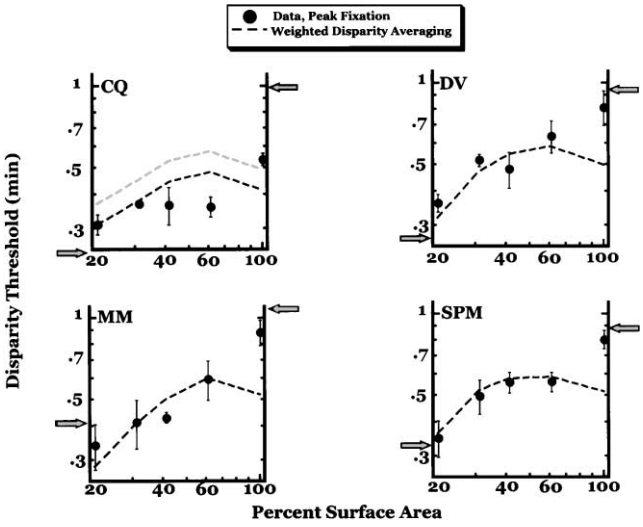

Fig. 6. Disparity thresholds as a function of the amount of surface presented for the partial surface stimuli. Data from 4 observers are shown. For each condition, the most curved stimulus (disparity = 11.2 min) was used. Data from the Peak Only condition (20% surface area) and the intact curved surface condition (100% surface area) are also plotted. As the percentage of surface increases, thresholds rise. For comparison, gray arrows along the left and right ordinates show each observers' disparity thresholds for a flat surface with 100% surface area. Thresholds for a flat surface in the fixation plane are shown at the left, and thresholds for a flat surface with 11.2 min of standing disparity are shown on the right. If observers were using just the peak of the surface to make a judgment, thresholds would be expected to be as low as those on the left ordinate. If observers were using just the back edges of the surface to make a judgment, thresholds would be expected to be as high as those on the right ordinate. The black dashed lines show threshold predictions for the weighted disparity averaging model for each observer. The gray dashed line shows observer CQ's unadjusted model predictions (see text). In general, the weighted disparity averaging model fits the data well for all points except the intact (100% surface) stimulus.

olds, consistent with the idea that the system is more precise with less disparity information. These outcomes strongly suggest that recovering a surface through multiple depths involves interaction among disparity detectors.

4. Discussion

Our results show that stereoacuity is impaired for curved luminance contour stimuli compared to stereoacuity for flat contour stimuli or curved random-dot stimuli. Our basic effect is that curved stimuli, which contain multiple disparities, have elevated thresholds relative to flat stimuli. We must qualify this basic outcome, however, because curved random-dot stimuli do not interfere with stereoacuity. Thus, we have two outcomes to explain: (1) why is sensitivity to curved luminance stimuli so poor compared to sensitivity for flat luminance stimuli? and (2) why is sensitivity to curved random-dot surfaces so much better than sensitivity to curved luminance surfaces?

4.1. Spatial frequency content

The luminance and random-dot stimuli appear, at first glance, to be composed of very different spatial frequency information. Dots are spatially restricted in both the vertical and horizontal dimensions and so their Fourier spatial frequency spectra in both dimensions will be uniform and very broad. The spectrum of an isolated luminance edge (i.e., a bright line) will differ from the spectrum of a dot only in the vertical dimension. The spectrum in the horizontal dimension, which is the dimension of interest with respect to horizontal disparity signals, is very similar for luminance edges and dots. Finally, the spectrum for a boundary contour in the horizontal dimension (specifically, a dark-to-light contour at the boundary of a surface of high luminance, as in our luminance surface stimuli) will differ from the spectrum of a dot or an isolated edge in that there will be less energy at the highest spatial frequencies. The results of Section 3.2 show that sensitivity to a single, isolated luminance contour is not different than sensitivity to the entire luminance surface. Thus, our data suggest that the relative differences in the amplitude of high spatial frequency information between dots and surface edges does not affect thresholds. We therefore consider it unlikely that differences in the spatial frequency content of the luminance and random-dot stimuli can account for the large difference in sensitivity. Further, a consideration of the spatial frequency content of our stimuli does not explain the difference in sensitivity between curved and flat luminance stimuli, which presumably have very similar spatial frequency content.

4.2. Disparity averaging

The results of the partial surface experiment strongly suggest that disparity detectors interact. One specific kind of disparity interaction that appears in the literature is disparity averaging (Parker & Yang, 1989; Rohaly & Wilson, 1994). Disparity averaging occurs when the visual system averages two different (and typically, spatially co-incident) disparities to arrive at a final estimate that is a linear combination of the two. If disparity averaging is occurring with our stimuli, then we should find the same thresholds for the luminance stimuli and random-dot stimuli, because both contain the same range of disparities. This was clearly not the case. It is possible, however, that disparity averaging does not occur in the same way for the two types of stimuli. Parker and Yang (1989), for example, used a random-dot test stimulus containing two disparities; observers matched the depth of this mixed-disparity test to a comparison containing the average of the two test disparities. They found that disparity averaging occurred over a limited range of disparities; the two components of the test stimulus could differ by a maximum of 3.3 min of disparity under some conditions and still yield disparity averaging. Rohaly and Wilson (1994) used similar logic with very different stimuli; their displays were composed of two gratings that differed in disparity and spatial frequency rather than two sets of disparate random dots. Rohaly and Wilson also found disparity averaging over a limited range of disparities, but argued that their data could not be explained by Parker and Yang's disparity averaging model. Thus, there is reason to suspect that disparity averaging does not occur in the same way for all stimuli. We were particularly interested in knowing whether we could predict thresholds for the partial luminance stimuli given a simple combination of independent disparity estimates. Our aim was to calculate the expected threshold given the multiple and discontinuous ranges of disparity in the partial surface stimuli of Fig. 5.

We generated threshold predictions by measuring each observer's thresholds for depth *intervals* as a function of distance from the horopter (McKee et al., 1990a; Siderov & Harwerth, 1995). For this measurement, we presented a flat surface in the fixation plane, and varied only the depth of the probe around the five standing disparities. The observer's task was to indicate whether the probe was in front of or behind the depth interval specified by the standing disparity. We thus obtained disparity thresholds for depth intervals as a function of disparity from the fixation plane, and we fit each observer's data with a linear function (R^2 varied from a low of 0.936 to a high of 0.992 among the observers). Thus, for any given disparity, we had an independent threshold estimate based on the depth interval task. Next, each of the partial surface stimuli was divided into 0.14 min disparity "bins". Each bin was assigned a threshold

estimate from the fitted linear function based on the disparity at the center of the bin. We then assigned a weight to each bin and averaged over all of the bins present in a particular stimulus. The weight for each bin was the inverse squared interval threshold; in other words, each bin's threshold estimate was weighted by the inverse variance. This procedure weights disparities close to fixation more heavily than disparities distant from fixation. Thus, our final threshold prediction for each stimulus was based on a weighted disparity averaging across the range of disparities present in each stimulus.

Fig. 6 shows the weighted disparity averaging prediction for each observer (black dashed line). Threshold predictions are low for the 20% (Peak Only) stimulus because this stimulus contains only a small range of disparities near fixation. The inverted U-shape of the disparity averaging predictions is the result of adding disparity from the back of the surface toward the peak. Disparities at the back of the surface are distant from the horopter, associated with high thresholds, and are not weighted very heavily. Thus, thresholds slowly increase from the 20% stimulus because disparities at the back of the surface are averaged with those at the peak. As more disparity is added from the back of the surface toward the peak, the trend changes directions. This is because the threshold associated with each added disparity decreases while its weight increases, generating the downward trend. Note that because predictions were based on individual threshold measurements, the model makes unique predictions for each observer. It is also important to note that observer CQ's thresholds in the interval task were elevated relative to his thresholds for curved surfaces in the main experiments. This may be the result of differential practice for this observer in the two types of task. Thus, CQ's raw model thresholds (shown in light gray in Fig. 6) were uniformly elevated relative to the data. We assumed that, with additional practice in the interval task, CQ's thresholds would decline without changing the shape of the function. Therefore, to generate a more informative prediction we pinned CQ's model thresholds to the datum at the Peak Only (20% surface area) condition. This procedure was unnecessary for the remaining three observers.

We have, in effect, constructed a very simple model of weighted disparity averaging based on a population of independent disparity detectors. Given these limitations, it is surprising how well the model accounts for most of our data. The thresholds for partial surface stimuli are, in general, well-described by weighted disparity averaging. The predictions fail, however, for the intact luminance surface. It is not obvious how modifications to the model (such as changing the weighting scheme or incorporating the size-disparity correlation by scaling the disparity bin width) could be made to account for the 100% surface luminance surface. Thus, we suggest that the visual system is calculating a simple weighted average of independent disparity signals when presented with the partial luminance surface stimuli, and that this process only occurs for incomplete surfaces. Indeed, Rohaly and Wilson (1994) have suggested that disparity averaging may be a useful strategy to generate the appearance of single, solid surfaces.

It is unclear why disparity averaging does not occur for the intact luminance surface. It seems likely, however, that an intact luminance boundary in depth generates more complex interactions between disparity detectors than simple averaging. It is also unclear why disparity averaging does not occur for random-dot stimuli in the same way as for luminance stimuli. Since performance for the random-dot surface resembles performance for the luminance, Peak Only stimulus, we speculate that the disparity information at the peak of the random-dot surface is selectively integrated in the horizontal dimension. There is recent evidence for horizontally-elongated disparity summation fields for surfaces that vary in depth about a horizontal axis (Tyler & Kontsevich, 2001). What remains unclear, however, is why this process does not influence thresholds for the intact luminance surface stimuli. Perhaps the absence of explicit disparity signals in the interior region of the luminance surface could account for this discrepancy.

To summarize and conclude, we return to the two questions posed at the beginning of Section 4. The superior performance for flat luminance stimuli relative to curved luminance stimuli can be explained by assuming that only disparity detectors tuned to the single disparity in flat stimuli contribute to the observed thresholds. We suggest that when incomplete luminance surfaces are presented, activity is generated in detectors tuned to many different disparities which undergo interactions. Thus, thresholds are generated by a weighted average of the sensitivities of the activated detectors. The intact curved surface, however, presumably generates more complex interactions between detectors tuned to various disparities. Here, thresholds may be a consequence of this more complex interaction. To answer the second question, we speculate that horizontally-elongated disparity summation fields allow random-dot stimuli to undergo selective integration. This integration restricts the disparity tuning of activated detectors and so mimics performance when only a single disparity is presented.

Although we cannot yet determine with certainty the mechanisms responsible for our data, we can say with greater certainty what mechanisms are not involved. The data cannot be accounted for by any straightforward application of the kind of disparity-energy processing assumed to occur in V1 (DeAngelis, Ohzawa, & Freeman, 1991). Further, the rules by which disparity signals are spatially combined are highly configuration- (and probably task-) dependent. Consider, for example, that McKee (1983) found that stereothresholds for two vertical lines increased dramatically when the lines were

joined by horizontal segments at the top and bottom to form a box, and that Fahle and Westheimer (1988) found depth judgments in a two-point depth discrimination task were adversely affected when an intervening point was added to the display. On the contrary, Kumar and Glaser (1992) found depth discrimination thresholds improved when intervening disparity features were added. There is growing evidence of stimulus configuration effects across many visual domains which cannot be accounted for by early linear cortical processing (Bonneh & Sagi, 1999; Olzak & Thomas, 1992; Verghese & Stone, 1997). This might be expected from a system that is arranged hierarchically—the results of neural interactions at one level of visual representation are passed onto the next stage, but not the pre-interaction input on which the results are based (Lennie, 1998). Our work points to a role for "higher-level" processing for disparity-defined contours and surfaces. There has been a great deal of recent neurophysiological work suggesting that areas V2, V3, and MT play an important role in the processing of disparity signals and depth contours (Adams & Zeki, 2001; Backus, Fleet, Parker, & Heefer, 2001; DeAngelis & Newsome, 1999; von der Heydt, Zhou & Friedman, 2000). There are also suggestions of 3D shape processing in inferior temporal cortex (Janssen, Vogels, Liu, & Orban, 2001; Janssen, Vogels, & Orban, 2000; Tanaka, Uka, Yoshiyama, Kato, & Fujita, 2001; Uka, Tanaka, Yoshiyama, Kato, & Fujita, 2000). It may be that some of the thresholds we observe in the current set of experiments are dependent on low-level disparity detectors, while other thresholds are generated by the activity of higher-level 3D contour or shape mechanisms.

Acknowledgements

We are very grateful to Doug Taylor for technical assistance and Mark Pettet for many helpful discussions. This work was supported by NEI EY12038 to PV, NASA NAG 9-1163 to PV, and NEI EY06644 to SPM.

References

Adams, D. L., & Zeki, S. (2001). Function organization of Macaque V3 for stereoscopic depth. *Journal of Neurophysiology, 86*, 2195–2203.

Badcock, D. R., & Schor, C. M. (1985). Depth increment detection function for individual spatial channels. *Journal of the Optical Society of America A, 2*, 1211–1216.

Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research, 39*, 1143–1170.

Backus, B. T., Fleet, D. J., Parker, A. J., & Heefer, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology, 86*, 2054–2068.

Berry, R. N. (1948). Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *Journal of Experimental Psychology, 38*, 708–721.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology, 211*, 599–622.

Bonneh, Y., & Sagi, D. (1999). Contrast integration across space. *Vision Research, 39*, 2597–2602.

Collett, T. S. (1985). Extrapolating and interpolating surfaces in depth. *Proceedings of the Royal Society of London B, 224*, 43–56.

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *Journal of Neuroscience, 19*, 1398–1415.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature, 352*, 156–159.

Fahle, M., & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research, 28*, 171–178.

Farell, B. (1998). Two-dimensional matches from one-dimensional stimulus components in human stereopsis. *Nature, 395*, 689–693.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review, 104*, 618–658.

Hess, R. F., Hayes, A., & Kingdom, A. A. (1997). Integrating contours within and through depth. *Vision Research, 37*, 691–696.

Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2001). Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *Journal of Neuroscience, 21*, 9419–9429.

Janssen, P., Vogels, R., & Orban, G. A. (2000). Three-dimensional shape coding in inferior temporal cortex. *Neuron, 27*, 385–397.

Kumar, T., & Glaser, D. A. (1992). Depth discrimination of a line is improved by adding other nearby lines. *Vision Research, 32*, 1667–1676.

Lennie, P. (1998). Single units and visual cortical organization. *Perception, 27*, 889–935.

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research, 23*, 191–198.

McKee, S. P., Levi, D. M., & Bowne, S. F. (1990a). The imprecision of stereopsis. *Vision Research, 30*, 1763–1779.

McKee, S. P., Welch, L., Taylor, D. G., & Bowne, S. F. (1990b). Finding the common bond: stereoacuity and the other hyperacuities. *Vision Research, 22*, 449–460.

Morgan, M. J., & Castet, E. (1997). The aperture problem in stereopsis. *Vision Research, 37*, 2737–2744.

Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America A, 43*, 906–913.

Olzak, L. A., & Thomas, J. P. (1992). Configural effects constrain fourier models of pattern discrimination. *Vision Research, 32*, 1885–1898.

Parker, A. J., & Yang, Y. (1989). Spatial properties of disparity pooling in human stereo vision. *Vision Research, 29*(11), 1525–1538.

Rohaly, A. M., & Wilson, H. R. (1994). Averaging across spatial scales. *Vision Research, 34*, 1315–1325.

Schumer, R. A., & Julesz, B. (1984). Binocular disparity modulation sensitivity to disparities offset from the plane of fixation. *Vision Research, 24*, 533–542.

Siderov, J., & Harwerth, R. S. (1995). Stereopsis, spatial frequency and retinal eccentricity. *Vision Research, 35*, 2329–2337.

Tanaka, H., Uka, T., Yoshiyama, K., Kato, M., & Fujita, I. (2001). Processing of shape defined by disparity in monkey inferior temporal cortex. *Journal of Neurophysiology, 85*, 735–744.

Tyler, C. W. (1974). Depth perception in disparity gratings. *Nature, 251*, 140–142.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research, 15*, 583–590.

Tyler, C. W., & Kontsevich, L. L. (2001). Stereoprocessing of cyclopean depth images: horizontally elongated summation fields. *Vision Research, 41*, 2235–2243.

Uka, T., Tanaka, H., Yoshiyama, K., Kato, M., & Fujita, I. (2000). Disparity selectivity of neurons in monkey inferior temporal cortex. *Journal of Neurophysiology, 84*, 120–132.

Verghese, P., & Stone, L. S. (1997). Spatial layout affects speed discrimination. *Vision Research, 37*, 397–406.

von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research, 40*, 1955–1967.

Westheimer, G., & McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Journal of the Optical Society of America A, 68*, 450–455.

Westheimer, G., & McKee, S. P. (1979). What prior uniocular processing in necessary for stereopsis? *Investigative Ophthalmology and Visual Sciences, 18*, 614–621.

Wurger, S. M., & Landy, M. S. (1989). Depth interpolation with sparse disparity cues. *Perception, 18*, 39–54.